ELSEVIER

Discrete Mathematics 196 (1999) 197–206

DISCRETE MATHEMATICS

On the degrees of irreducible factors of polynomials over a finite field

Arnold Knopfmacher*

Department of Computational & Applied Mathematics, University of the Witwatersrand PO Wits, 2050 South Africa

Received 4 September 96; revised 10 February 98; accepted 16 March 98

Abstract

Let $\tilde{\mathbb{F}}_q[X]$ denote the multiplicative semigroups of monic polynomials in one indeterminate X, over a finite field $\mathbb{F}_q$. We determine for each fixed q, the asymptotic average number of distinct degrees of the irreducible factors of a polynomial of degree n in $\tilde{\mathbb{F}}_q[X]$. The results are of relevance to various polynomial factorization algorithms.

1. Introduction

Let $\tilde{\mathbb{F}}_q[X]$ denote the multiplicative semigroup of monic polynomials in one indeterminate X over a finite field with q elements. Many deterministic as well as probabilistic factorization algorithms for polynomials in $\tilde{\mathbb{F}}_q[X]$ require that a distinct degree factorization of the polynomial be performed as an initial step. This method factors a polynomial f into factors f_d such that f_d is the product of all the irreducible factors of degree d. See, e.g. [10, pp. 429–431] and more recently [1–3,7,11]. Further references can be found in Shparlinski [12, Ch. 1]. If a given polynomial f in $\tilde{\mathbb{F}}_q[X]$ has k irreducible factors of r different degrees, then at most r polynomials f_d require factorization (if not already irreducible) and a further $k-r$ irreducible factors need to be extracted in order to obtain the complete factorization of f . It is therefore of interest to determine the average value of r and to compare this to the known average values for the number of irreducible factors, or of distinct irreducible factors of such polynomials. In the sequel, the value r will be termed the number of *irreducible degrees* of a polynomial.

Our main result below illustrates the perhaps surprising fact that on average the distinct degree factorization of f in $\tilde{\mathbb{F}}_q[X]$ is already rather close to the full factorization of f.

* E-mail: arnoldk@gauss.cam.wits.ac.za.

0012-365X/99/$ – see front matter
PII: S0012-365X(98)00174-5

2. Results

As is well known there are q^n monic polynomials of degree n in $\mathbb{F}_q[X]$ and

$$\pi(n) \equiv \pi(n,q) = \frac{1}{n}\sum_{d|n} \mu(d) q^{n/d} \tag{1}$$

monic irreducible polynomials of degree n, where $\mu(\cdot)$ denotes the Möbius function.

The principal result of the paper is the following:

Theorem 1. *The number r of irreducible degrees of a polynomial of degree n in $\tilde{\mathbb{F}}_q[X]$ has mean value*

$$\log n + \gamma + c_1(q) - A(q) + \mathrm{O}\left(\frac{1}{n}\right) \quad \text{as } n \to \infty,$$

where γ is Euler's constant,

$$c_1(q) = \sum_{r=1}^{\infty} (\pi(r) - q^r/r) q^{-r} < 0$$

and

$$A(q) = \sum_{m=1}^{\infty} \left((1 - q^{-m})^{\pi(m)} - 1 + \frac{\pi(m)}{q^m} \right). \tag{2}$$

In addition almost all polynomials of degree n have approximately $\log n$ *irreducible degrees (i.e. the* normal order *is* $\log n$*).*

Remarks. It is well known (see [8] and the references therein) that the average number $\omega(n)$ of distinct irreducible factors of f in $\tilde{\mathbb{F}}_q[X]$ of degree n is

$$\log n + \gamma + c_1(q) + \mathrm{O}\left(\frac{1}{n}\right), \quad n \to \infty.$$

From this we deduce

Corollary 1. *The average number of distinct irreducible factors exceeds the average number of irreducible degrees of a polynomial of degree n in $\tilde{\mathbb{F}}_q[X]$ by an amount that approaches the constant*

$$A(q) = \sum_{m=1}^{\infty} \left((1 - q^{-m})^{\pi(m)} - 1 + \frac{\pi(m)}{q^m} \right)$$

as n approaches infinity.

In addition (see [8]) the corresponding average $\Omega(n)$ for the total number of irreducible factors (including multiplicity)is

$$\omega(n) + c_2(q) + O\left(\frac{1}{n}\right),$$

where

$$0 < c_2(q) = \sum_{r=1}^{\infty} \frac{\pi(r)}{q^r(q^r - 1)}.$$

Hence also as $n \to \infty$, the total number of irreducible factors of f of degree n in $\mathbb{F}_q[X]$ exceeds the number of irreducible degrees on average by

$$A(q) + c_2(q).$$

A table of approximate values for these constants for small values of q is given below.

q	$c_1(q)$	$c_2(q)$	$A(q)$	$A(q)+c_2(q)$
2	−0.4522339	1.1417854	0.3717492	1.5135347
3	−0.2439834	0.5569674	0.4939232	1.0508907
4	−0.1660084	0.3644586	0.5436804	0.9081390
5	−0.1254551	0.2697086	0.5705894	0.8402980
7	−0.0840230	0.1766655	0.5989254	0.7755909
8	−0.0720594	0.1505106	0.6072862	0.7577968
9	−0.0630602	0.1310491	0.6136335	0.7446827
∞	0	0	0.6598155	0.6598155

A related problem of interest is to determine the proportion $L_n(q)$ of polynomials of degree n in $\tilde{\mathbb{F}}_q[X]$ for which the distinct degree factorization is the full factorization of f. It is shown in [9] that

$$\lim_{n\to\infty} L_n(q) \equiv L(q) = \prod_{m=1}^{\infty} \left(1 + \frac{\pi(m)}{q^m}\right) \exp(-1/m).$$

Furthermore, $L(q) > 0.39$ for $q \geqslant 2$ and $L(q) \to e^{-\gamma}$ as $q \to \infty$. Corollary 1 now shows that the remaining set of polynomials of degree n in $\tilde{\mathbb{F}}_q[X]$ are also "close" to having a distinct degree factorization on average.

Further results related to the above can be found in the recent paper of Flajolet et al. [4].

3. Analysis

Let $\partial f(X)$ denote the degree of $f(X)$ in $\tilde{\mathbb{F}}_q[X]$.

Let P_n denote the set of monic irreducible polynomials of degree n in $\mathbb{F}_q[X]$. Thus $|P_n| = \pi(n)$.

Let $\rho(n,k)$ denote the number of polynomials in $\mathbb{F}_q[X]$ of degree n with exactly k different degrees of irreducible factors. As is familiar the generating function for the number of monic polynomials of degree n in $\mathbb{F}_q[X]$ is

$$\frac{1}{1-qx} = \prod_{m=1}^{\infty} \prod_{p \in P_m} (1 + x^{\partial p} + x^{2\partial p} + x^{3\partial p} + \cdots)$$

$$= \prod_{m=1}^{\infty} (1 - x^m)^{-\pi(m)}. \tag{3}$$

To obtain the bivariate generating function for $\rho(n,k)$ we tag with y the product over $p \in P_m$ for each m, with the exception of the case that no irreducible factor of degree m appears. This leads to

$$\rho(x,y) := \sum_{n,k} \rho(n,k) x^n y^k$$

$$= \prod_{m=1}^{\infty} \left(1 + y\left(\prod_{p \in P_m} (1 + x^{\partial p} + x^{2\partial p} + x^{3\partial p} + \cdots) - 1\right)\right)$$

$$= \prod_{m=1}^{\infty} (1 + y((1 - x^m)^{-\pi(m)} - 1)). \tag{4}$$

Alternatively, we may deduce (4) from the general formula obtained by Wilf [13, Eq. (12)] for objects of size n in combinatorial structures whose parts have k different sizes.

Now, let $\bar{\rho}(n)$ denote the average number of irreducible degrees of a polynomial of degree n in $\tilde{\mathbb{F}}_q[X]$. We obtain by differentiation

$$\sum_{n=1}^{\infty} \bar{\rho}(n) q^n x^n = \frac{\partial}{\partial y} \rho(x,y)|_{y=1}$$

$$= \prod_{m=1}^{\infty} (1 - x^m)^{-\pi(m)} \sum_{m \geqslant 1} (1 - x^m)^{\pi(m)} ((1 - x^m)^{-\pi(m)} - 1)$$

$$= \frac{1}{1-qx} \sum_{m \geqslant 1} (1 - (1 - x^m)^{\pi(m)}).$$

Thus,

$$\bar{\rho}(n) = q^{-n}[x^n] \sum_{m \geqslant 1} \left(\frac{1 - (1 - x^m)^{\pi(m)}}{1 - qx}\right). \tag{5}$$

A point of particular interest concerning this generating function is that it admits the circle $|x| = \frac{1}{q}$ as a natural boundary. Consequently, the technique of *singularity analysis* as developed recently by Flajolet and Odlyzko [5] cannot be applied in this case. Instead our analysis will be based on asymptotic estimation of sums.

We have

$$(1-x^m)^{\pi(m)}(1-qx)^{-1}=\left\{\sum_{r=0}^{\pi(m)}\binom{\pi(m)}{r}(-x^m)^r\right\}\left\{\sum_{r=0}^{\infty}q^r x^r\right\}.$$

Hence,

$$\begin{aligned}\bar{\rho}(n)&=\sum_{m=1}^{n}\left(1-\sum_{r=0}^{\lfloor n/m\rfloor}(-1)^r\binom{\pi(m)}{r}q^{-mr}\right)\\&=\sum_{m=1}^{n}\sum_{r=1}^{\lfloor n/m\rfloor}(-1)^{r+1}\binom{\pi(m)}{r}q^{-mr}\\&=\sum_{m=1}^{n}\frac{\pi(m)}{q^m}-\sum_{m=1}^{\lfloor n/2\rfloor}\sum_{r=2}^{\lfloor n/m\rfloor}(-1)^r\binom{\pi(m)}{r}q^{-mr}.\end{aligned}\tag{6}$$

As shown for example in [8]

$$\sum_{m=1}^{n}\frac{\pi(m)}{q^m}=\omega(n)=\log n+\gamma+c_1(q)+\mathrm{O}\left(\frac{1}{n}\right),\tag{7}$$

where $\omega(n)$ is the average value of the number of distinct irreducible factors of f of degree n in $\tilde{\mathbb{F}}_q[X]$. It remains therefore to estimate the double sum.

It will be useful later to note that the exact formula (1) for $\pi(n)$ provides the bounds (see e.g. [9, p. 2241]),

$$q^n-2q^{n/2}<n\pi(n)<q^n.\tag{8}$$

Firstly, for $n\geqslant m\pi(m)$ we have

$$\sum_{r=2}^{\lfloor n/m\rfloor}(-1)^r\binom{\pi(m)}{r}q^{-mr}=(1-q^{-m})^{\pi(m)}-1+\frac{\pi(m)}{q^m}.$$

Next for $n<m\pi(m)$,

$$\begin{aligned}\sum_{r=2}^{\lfloor n/m\rfloor}(-1)^r\binom{\pi(m)}{r}q^{-mr}&=(1-q^{-m})^{\pi(m)}-1+\frac{\pi(m)}{q^m}\\&\quad+\mathrm{O}\left(\sum_{r=\lfloor n/m\rfloor+1}^{\pi(m)}\binom{\pi(m)}{r}q^{-mr}\right).\end{aligned}$$

It follows from (8) that $m\geqslant\log_q n$ in this case, and

$$\sum_{r=\lfloor n/m\rfloor+1}^{\pi(m)}\binom{\pi(m)}{r}q^{-mr}<\sum_{r=\lfloor n/m\rfloor+1}^{\infty}\frac{1}{m^r r!}<\sum_{r=0}^{\infty}\frac{1}{m^{n/m}r!}=\frac{e}{m^{n/m}}.$$

Now $m^{n/m}$ is a decreasing function of m for $\log_q n \leqslant m \leqslant n/2$ and so achieves a minimum at $m = n/2$. Consequently,

$$\sum_{r=\lfloor n/m\rfloor+1}^{\infty} \frac{1}{m^r r!} = \mathrm{O}\left(\frac{1}{n^2}\right).$$

Hence,

$$\begin{aligned}\sum_{m=1}^{\lfloor n/2\rfloor}\sum_{r=2}^{\lfloor n/m\rfloor}(-1)^r\binom{\pi(m)}{r}q^{-mr} &= \sum_{m=1}^{\lfloor n/2\rfloor}\left((1-q^{-m})^{\pi(m)} - 1 + \frac{\pi(m)}{q^m}\right) + \mathrm{O}\left(\frac{1}{n}\right)\\ &= \sum_{m=1}^{\infty}\left((1-q^{-m})^{\pi(m)} - 1 + \frac{\pi(m)}{q^m}\right) + \mathrm{O}\left(\frac{1}{n}\right)\\ &\quad + \mathrm{O}\left(\sum_{m=\lfloor n/2\rfloor+1}^{\infty}\left((1-q^{-m})^{\pi(m)} - 1 + \frac{\pi(m)}{q^m}\right)\right).\end{aligned}$$

By (8) the last sum is

$$\mathrm{O}\left(\sum_{m\geqslant n/2}\left(\mathrm{e}^{-1/m} - 1 + \frac{1}{m}\right)\right) = \mathrm{O}\left(\sum_{m\geqslant n/2}\frac{1}{m^2}\right) = \mathrm{O}\left(\frac{1}{n}\right).$$

Hence,

$$\omega(n) - \bar{\rho}(n) = A(q) + \mathrm{O}\left(\frac{1}{n}\right),$$

where

$$A(q) = \sum_{m=1}^{\infty}\left((1-q^{-m})^{\pi(m)} - 1 + \frac{\pi(m)}{q^m}\right)$$

and the first part of Theorem 1 follows.

Note that if we let $q \to \infty$ then $\pi(m)/q^m \to 1/m$ for each m, and

$$\begin{aligned}A(q) \to A &= \sum_{m=1}^{\infty}(\mathrm{e}^{-1/m} - 1 + 1/m)\\ &= \sum_{m=1}^{\infty}\sum_{j=2}^{\infty}\frac{(-1)^j}{m^j j!} = \sum_{j=2}^{\infty}\frac{(-1)^j}{j!}\zeta(j) = 0.6598152\cdots,\end{aligned}$$

where $\zeta(j)$ denotes the Riemann zeta function.

This same constant gives the excess as $n \to \infty$ of the average number of cycles in a random n-permutation over the average number of distinct cycle lengths [13].

In addition since $[1-(1-x^m)^{\pi(m)}]/(1-qx)$ is the generating function for the number of polynomials of degree n in $\tilde{\mathbb{F}}_q[X]$ with at least one irreducible factor of degree m, it follows from [5] that the average value $\bar{\rho}(n)$ *is equal to the sum of the proportions of monic polynomials of degree n having at least one irreducible factor of degree m*, $m=1,2,\ldots n$.

By contrast the expression

$$\omega(n)=\sum_{m=1}^{n}\frac{\pi(m)}{q^m}$$

implies that $\omega(n)$ *is the sum of proportions of polynomials of degree m that are irreducible*, $m=1,2,\ldots,n$.

Let $\rho(f)$ be the number of irreducible degrees of a given f in $\tilde{\mathbb{F}}_q[X]$. The variance of the number of different degrees of irreducible factors for polynomials of degree n in $\tilde{\mathbb{F}}_q[X]$ is $q^{-n}\sum_{\partial f=n}(\rho(f)-\bar{\rho}(n))^2$ which is given by the formula

$$q^{-n}[x^n]\left\{\frac{\partial^2}{\partial y^2}\rho(x,y)|_{y=1}\right\}+\bar{\rho}(n)-\bar{\rho}(n)^2. \tag{9}$$

To prove the result about the normal order of $\rho(f)$ it is sufficient to show that

$$[x^n]q^{-n}\frac{\partial^2}{\partial y^2}\rho(x,y)|_{y=1}\sim\bar{\rho}^2(n)\sim\log^2 n \tag{10}$$

since then the claimed result follows from Chebyshev's inequality.

Now,

$$\frac{\partial^2}{\partial y^2}\rho(x,y)=\frac{\partial}{\partial y}\left\{\rho(x,y)\sum_{m=1}^{\infty}\frac{(1-x^m)^{-\pi(m)}-1}{1+y((1-x^m)^{-\pi(m)}-1)}\right\}.$$

Hence,

$$\frac{\partial^2}{\partial y^2}\rho(x,y)|_{y=1}$$
$$=\frac{1}{1-qx}\left\{\left\{\sum_{m=1}^{\infty}(1-(1-x^m)^{\pi(m)})\right\}^2-\sum_{m=1}^{\infty}(1-(1-x^m)^{\pi(m)})^2\right\}.$$

Firstly, we consider

$$Q_n:=[x^n]\frac{1}{1-qx}\left\{\sum_{m=1}^{\infty}(1-(1-x^m)^{\pi(m)})\right\}^2$$
$$=[x^n]\left\{\sum_{m=1}^{\infty}q^m\bar{\rho}(m)x^m\right\}\left\{\sum_{m=1}^{\infty}(1-(1-x^m)^{\pi(m)})\right\}.$$

Let

$$\sum_{n=1}^{\infty} a_n q^n x^n := \sum_{m=1}^{\infty} (1-(1-x^m)^{\pi(m)})$$
$$= \sum_{m=1}^{\infty} \sum_{r=1}^{\pi(m)} \binom{\pi(m)}{r} (-1)^{r+1} x^{mr}$$
$$= \sum_{n=1}^{\infty} \sum_{mr=n} \binom{\pi(m)}{r} (-1)^{r+1} x^n.$$

Therefore,

$$a_n q^n = \pi(n) + \sum_{\substack{d|n \\ d>1}} \binom{\pi(n/d)}{d} (-1)^{d+1}.$$

Now,

$$\left| \sum_{\substack{d|n \\ d>1}} \binom{\pi(n/d)}{d} (-1)^{d+1} \right| \leqslant \sum_{d=2}^{n} \frac{\pi^d(n/d)}{d!}$$
$$\leqslant q^n \sum_{d=2}^{n} \frac{1}{d!} \left(\frac{d}{n}\right)^2 \quad \text{(by (8))}$$
$$< \frac{q^n}{n^2} \sum_{d=2}^{\infty} \frac{2}{(d-2)!} = O\left(\frac{q^n}{n^2}\right).$$

Hence by (8),

$$a_n = \frac{1}{n} + O\left(\frac{1}{n^2}\right). \tag{11}$$

Now Theorem 1 implies that with an explicit constant $c(q) = \gamma + c_1(q) - A(q)$,

$$\bar{\rho}(n) = \log n + c(q) + O\left(\frac{1}{n}\right) \quad \text{as } n \to \infty. \tag{12}$$

Thus,

$$Q_n = \sum_{i=1}^{n-1} q^i a_i q^{n-i} \bar{\rho}(n-i)$$
$$= q^n \sum_{i=1}^{n-1} \left(\frac{1}{i} + O\left(\frac{1}{i^2}\right)\right) \left(\log(n-i) + c(q) + O\left(\frac{1}{n-i}\right)\right)$$
$$= q^n \left\{ \sum_{i=1}^{n-1} \frac{\log(n-i)}{i} + O\left(\sum_{i=1}^{n-1} \frac{1}{i}\right) + O\left(\sum_{i=1}^{n-1} \frac{\log(n-i)}{i^2}\right) \right\}.$$

As shown for example in [8, p. 115],

$$\sum_{i=1}^{n-1} \frac{\log(n-i)}{i} = \log^2 n + O(\log n).$$

It follows that

$$Q_n = q^n\{\log^2 n + O(\log n)\}. \tag{13}$$

Now, consider

$$R_n := [x^n]\frac{1}{1-qx}\sum_{m=1}^{\infty}\left(1-(1-x^m)^{\pi(m)}\right)^2 := [x^n]\frac{B(x)}{1-qx}.$$

For $|x| \leqslant 1/q$ by (8),

$$|1-(1-x^m)^{\pi(m)}|^2 \leqslant \left(\left(1+\frac{|qx|^m}{q^m}\right)^{q^m/m} - 1\right)^2$$
$$\leqslant (e^{|qx|^m/m} - 1)^2 = O\left(\frac{|qx|^{2m}}{m^2}\right), \quad m = 1,2,3,\ldots.$$

Hence $B(x)$ is absolutely convergent for $|x| \leqslant 1/q$ and it follows that

$$R_n = O(q^n). \tag{14}$$

From (13) and (14), (10) follows.

We note that with some additional work (13) can be improved to

$$Q_n = q^n(\log^2 n + 2c(q)\log n + O(1)), \tag{15}$$

with $c(q)$ as previously defined.

Then by substituting (12), (14) and (15) into (9) we obtain for the variance the estimate $\log n + O(1)$, as $n \to \infty$.

Finally, we remark that results of Flajolet and Soria [6] imply that the number of irreducible factors of polynomials of degree n in $\tilde{\mathbb{F}}_q[x]$ has a limiting normal distribution. The corresponding limit distribution for the number of irreducible degrees is the subject of a further work currently in progress between the author and P. Grabner.

Acknowledgements

The author wishes to thank P. Grabner, D.S. Lubinsky, R. Warlimont and the referees for helpful comments concerning an earlier draft of the manuscript.

References

[1] E. Bach, V. Shoup, Factoring polynomials using fewer random bits, J. Symbolic Comput. 9 (1990) 229–239.
[2] M. Ben-Or, Probabilistic algorithms in finite fields, Proc. 22nd Annual Symp. Foundations of Computer Science, 1981, pp. 394–398.
[3] D. Cantor, H. Zassenhaus, A new algorithm for factoring polynomials over finite fields, Math. Comput. 36 (1981) 587–592.
[4] P. Flajolet, X. Gourdon, D. Panario, Random polynomials and polynomial factorization, Lecture Notes in Computer Science, vol. 1099 (1996) 232–243.
[5] P. Flajolet, A. Odlyzko, Singularity Analysis of generating functions, SIAM J. Discrete Math. 3 (1990) 216–240.
[6] P. Flajolet, M. Soria, Gaussian limiting distributions for the number of components in combinatorial structures, J.C.T. Series A 53 (1990) 165–182.
[7] J. von zur Gathen, V. Shoup, Computing Frobenius maps and factoring polynomials, Comput. Complexity 2 (1992) 187–224.
[8] A. Knopfmacher, J. Knopfmacher, Counting irreducible factors of polynomials over a finite field, Discrete Math. 112 (1993) 103–118.
[9] A. Knopfmacher, R. Warlimont, Distinct degree factorizations for polynomials over a finite field, Trans. Amer. Math. Soc. 347 (1995) 2235–2243.
[10] D.E. Knuth, The art of computer programming, vol. 2, 2nd ed., Addison-Wesley, Reading, MA, 1981.
[11] V. Shoup, On the deterministic complexity of factoring polynomials over finite fields, Inform. Process. Lett. 33 (1990) 261–267.
[12] I.E. Shparlinski, Computational and algorithmic problems in finite fields, Kluwer, Dordrecht, 1992.
[13] H.S. Wilf, Three problems in combinatorial asymptotics, J.C.T. Ser. A 35 (1983) 199–207.